

# CONSISTENT RUN SELECTION FOR INDEPENDENT COMPONENT ANALYSIS: APPLICATION TO FMRI ANALYSIS

Qunfang Long<sup>1,\*</sup>, Chunying Jia<sup>1,\*</sup>, Zois Boukouvalas<sup>2</sup>, Ben Gabrielson<sup>1</sup>, Darren Emge<sup>1</sup>, and Tülay Adalı<sup>1</sup>

<sup>1</sup>Dept. of CSEE, University of Maryland, Baltimore County, MD, 21250

<sup>2</sup>Dept. of ENME, University of Maryland, College Park, MD, 20742

## ABSTRACT

Independent component analysis (ICA) has found wide application in a variety of areas, and analysis of functional magnetic resonance imaging (fMRI) data has been a particularly fruitful one. Maximum likelihood provides a natural formulation for ICA and allows one to take into account multiple statistical properties of the data—forms of diversity. While use of multiple types of diversity allows for additional flexibility, it comes at a cost, leading to high variability in the solution space. In this paper, using simulated as well as fMRI-like data, we provide insight into the trade-offs between estimation accuracy and algorithmic consistency with or without deviations from the assumed model and assumptions such as the statistical independence. Additionally, we propose a new metric, cross inter-symbol interference, to quantify the consistency of an algorithm across different runs, and demonstrate its desirable performance for selecting consistent run compared to other metrics used for the task.

**Index Terms**— Consistent run selection, independent component analysis, fMRI analysis

## 1. INTRODUCTION

Independent component analysis (ICA) is a powerful method for blind source separation (BSS) that extracts the underlying sources—latent variables—through the assumption of statistical independence, and by exploiting various statistical properties of these sources, *i.e.*, forms of diversity [1]. Due to the limited assumptions placed on the data, ICA has been applied to many biomedical imaging applications, such as the analysis of functional magnetic resonance imaging (fMRI) [1, 2] or electroencephalograph (EEG) [3, 4] data.

The maximum likelihood (ML) principle provides a natural umbrella for ICA under which most ICA algorithms can be derived [1, 5, 6], also provides a number of advantages such as asymptotic optimality as well as enabling use of all relevant forms of diversity through appropriately selected models for the source probability density function (PDF) [6]. Parameter estimation in ICA is generally performed in two stages. The first stage consists of the estimation of the source PDF while the second stage consists of maximization of source independence through iterative estimation of the demixing matrix. As additional forms of diversity are incorporated into the algorithm, it becomes more flexible. However, this flexibility comes at the expense of a more complicated cost function surface, affecting robustness of iterative algorithms, *e.g.*, to initializations, as in the bias-variance trade-off in estimation theory. For example, the fixed nonlinearity of Infomax [7] results in a smaller solution space than ICA by entropy bound minimization (ICA-EBM) [8]—which uses a flexible nonlinearity—and thus yields more stable solutions. But unless fixed nonlinearities are justified through prior knowledge of the sources, methods that use such nonlinearities will incur a bias in proportion to the difference between modeled and true PDFs that can be quantified with Kullback-Leibler (KL)-divergence [9].

\*Qunfang Long and Chunying Jia contributed equally to this work.

\*\*This work was supported in part by the grants NSF-CCF 1618551 and NSF 1631838.

In applications to real world data, a solution that has been proposed is to perform a number of independent runs and select a single run among those as the final estimate. Most notably the ICASSO approach includes visualization tools to study the consistency of an algorithm and the solution space and selects a set of representative components among those of multiple runs of an algorithm [10]. In an application like medical image analysis obviously selection of the most representative run is critical as each independent component describes a brain network and is considered to be a “fingerprint” [11] or a biomarker of disease [12]. ICASSO and later its modified version m-ICASSO [13] to select a single run rather than just centroids as in ICASSO, and more recently minimum spanning tree (MST) [14] have been proposed. These are included in the group ICA of fMRI toolbox (GIFT) [15] that has been widely used for ICA of fMRI data [16].

As noted by the nonparametric, prediction, activation, influence, reproducibility, and resampling (NPAIRS) framework in neuroimaging [17, 18], both accuracy and prediction are key considerations in algorithm performance. Thus, we first address the question of whether how reliable consistent runs are in terms of accurate estimation performance, and show that as one would expect, when all model assumptions are well satisfied, the most consistent run is also very likely to be the most accurate run. But if model mismatch occurs, it is more likely for the most accurate run to be an outlier that is not easily reproducible.

We study the trade-offs between estimation accuracy and algorithmic consistency using fMRI-like data as well as simulated data drawn from the generalized Gaussian distribution (GGD). Three ICA algorithms, Infomax, ICA-EBM and SparseICA-EBM [19], are used. Infomax, although not flexible like ICA-EBM, is the most widely used algorithm in fMRI analysis [1]. It uses a fixed nonlinearity, which is a good match for fMRI sources of interest and is also the default option in GIFT. ICA-EBM, on the other hand, is flexible in terms of matching the underlying source PDF and has been shown to produce desirable results for fMRI analysis [20], hence can better maximize independence when the sources are truly independent. SparseICA-EBM, an extension of ICA-EBM, incorporates sparsity as an additional form of diversity and thus can obtain meaningful decompositions even when the assumption of independence is not true in a strict sense. We study the performance of m-ICASSO and MST, and introduce a new practical metric, cross inter-symbol interference (cross-ISI) to evaluate consistency and to select a single run in real world applications where the ground truth is not available. We show its desirable performance and computational efficiency enabling significantly larger number of runs than m-ICASSO and MST.

## 2. METHODS

### 2.1. ICA

Let  $N$  statistically independent latent sources  $\mathbf{s}(v) = [s_1(v), \dots, s_N(v)]^T$  be mixed through an unknown invertible matrix  $\mathbf{A} \in \mathbb{R}^{N \times N}$  so that the mixtures  $\mathbf{x}(v) = [x_1(v), \dots, x_N(v)]^T$ , are

obtained through the linear model

$$\mathbf{x}(v) = \mathbf{A}\mathbf{s}(v), v = 1, \dots, V$$

where  $v$  denotes the discrete sample index and  $(\cdot)^\top$  the transpose. The goal of ICA is to estimate a demixing matrix  $\mathbf{W} \in \mathbb{R}^{N \times N}$  such that to produce source estimates  $\mathbf{y} = [y_1, \dots, y_N]^\top$  given by  $\mathbf{y}(v) = \mathbf{W}(v)\mathbf{x}(v)$ . In what follows we drop index  $v$  for simplicity.

ICA can be achieved by minimizing the mutual information (MI) among the underlying sources. The MI cost function defined as the KL-divergence between the joint source density and the product of the marginal estimated source densities is given by

$$\mathcal{J}(\mathbf{W}) = \sum_{n=1}^N H(y_n) - \log |\det(\mathbf{W})| - H(\mathbf{x}), \quad (1)$$

where  $H(\cdot)$  is the differential entropy and  $H(\mathbf{x})$  is a term independent of  $\mathbf{W}$  and can be ignored. Minimization of MI of the estimated sources can be shown to be equivalent to the ML function when the model PDF and the true PDF coincide for each source [1]. Let  $\hat{p}_{s_n}(y_n)$  denote the estimated PDF of the  $n$ th source. Then the differential entropy is given by

$$H(y_n) = -f(p_{s_n}(y_n), \hat{p}_{s_n}(y_n)) - E\{\log \hat{p}_{s_n}(y_n)\}, \quad (2)$$

where  $f(p_{s_n}(y_n), \hat{p}_{s_n}(y_n))$  denotes the KL-divergence between the  $n$ th estimated and the true source PDF. Once the model deviates from the true PDF, a bias is introduced in the estimate of the demixing matrix.

The differences in the assumed latent source models in the ML formulation lead to differences in the separation performance of ICA algorithms. Infomax, the most widely used algorithm for fMRI analysis, uses a fixed nonlinearity [7] which is a good match for a specific super-Gaussian distribution. Due to use of fixed nonlinearity Infomax results in a small solution space. However, its performance suffers when the true density deviates from this assumed model. ICA-EBM provides flexible density matching through the use of four measuring functions based on the maximum entropy principle [8] leading to smaller bias, *i.e.*, making it more likely to have  $f(p_{s_n}(y_n), \hat{p}_{s_n}(y_n)) \rightarrow 0$ , but at the expense of a relatively large solution space. Another very recent algorithm, SparseICA-EBM algorithm, inherits the advantage of ICA-EBM, namely its flexibility with enhanced performance due to the exploitation of the sparsity of the underlying sources through a regularization parameter  $\lambda$  that introduces sparsity into source estimates [19, 21]. The assumption of ICA that the underlying sources are statistically independent can be too restrictive an assumption in practice [19]. Incorporating relevant prior information, *e.g.*, sparsity, can relax the independence assumption, resulting in better match to the underlying problem. The introduced regularization term also helps to shrink the solution space for ICA-EBM, especially with increasing  $\lambda$ .

## 2.2. Algorithmic consistency

Most ICA algorithms derived under the ML umbrella are of iterative nature, and especially with the use of more flexible density models, solutions might be significantly different depending on the initialization. Thus selecting the solution for further analysis is an important problem, which has been addressed for the application of ICA to analysis of fMRI data. There have been two solutions proposed for this task, m-ICASSO [13] and MST [14]. In both cases, ICA algorithm is run  $R$  times using different initializations and in the second stage, different procedures are employed to identify the final run that will be used as the most consistent one and for further analysis, *e.g.*, to determine biomarkers of disease between patient and healthy groups of subjects.

### 2.2.1. M-ICASSO and MST

To study the reliability of estimates from multiple ICA runs, ICASSO is introduced as an explorative visualization method [10]. This method groups all  $N \times R$  estimates into clusters where  $N$  is the total number of components estimated in each of  $R$  runs. The reliability of an estimate is quantified by the quality, say compactness, of the cluster to which it is assigned. In ICASSO, the centroid estimates of all given clusters are put together and used for post analysis. This breaks the connection to the linear mixing model assumption in ICA since the centroids have not been generated from the same run and thus the uniqueness of the decomposition is violated. This is not practical for applications such as the analysis of fMRI data where access to both spatial and temporal estimated components is important. Therefore, in [13], ICASSO is modified to select a given run based on  $Q$ -index, which is a measure of the closeness of a single run to the centroids. We refer to this method as modified ICASSO (m-ICASSO). To calculate the  $Q$ -index, m-ICASSO takes into account the clusters that are identified according to their size and compactness using a given threshold. This makes the procedure somewhat subjective. Additionally, m-ICASSO does not allow for all the estimates of a given run to be included in the computation of the  $Q$ -index. Because if two or more components of a run are grouped into the same cluster, only the component that is most similar to the cluster's centroid is included in the computation. This leads loss of information.

In contrast, MST identifies a central run according to the alignment cost of all the components [14, 22]. Components of all the other runs are aligned with respect to those of the central run with the minimum cost. Afterwards, a one-sample  $t$ -test is performed for each type of component across all  $R$  runs, producing a  $T$ -map. These  $T$ -maps are good delineation of the functional networks thus are especially practically useful for fMRI analysis. The final run is selected as the one whose components are the most correlated to the  $T$ -maps. MST successfully addresses the previous issues. However, the use of  $T$ -maps limits its application to only fMRI and similar applications.

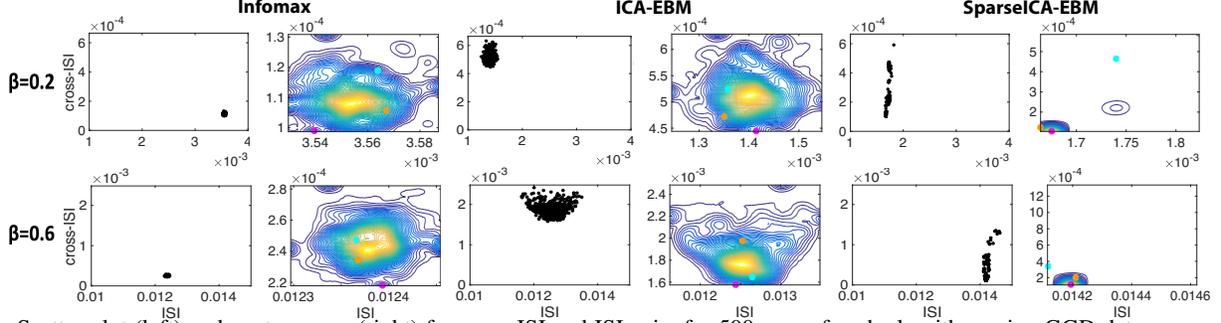
The consistency of the final run selected by m-ICASSO and MST has not been well studied, and both procedures as we highlight in the results section are computationally costly.

### 2.2.2. Cross-ISI

ISI is a frequently used global metric for performance evaluation when the ground truth is available. It is defined as:

$$\text{ISI}(\mathbf{G}) = \frac{1}{2N(N-1)} \cdot \sum_{n=1}^N \left( \sum_{m=1}^N \frac{\|g_{nm}\|}{\max_k \|g_{nk}\|} - 1 \right) + \frac{1}{2N(N-1)} \cdot \sum_{m=1}^N \left( \sum_{n=1}^N \frac{\|g_{nm}\|}{\max_k \|g_{km}\|} - 1 \right) \quad (3)$$

where  $\mathbf{G} = \mathbf{A}\mathbf{W}$  with elements denoted as  $g_{nm}$ , where  $\mathbf{A}$  is the true mixing matrix and  $\mathbf{W}$  is the estimated demixing matrix. If  $\mathbf{W}$  is perfectly estimated,  $\mathbf{G}$  is identity subject to permutation and scaling ambiguities, thus yielding zero ISI which indicating perfect separation. Therefore, the smaller the ISI, the closer the estimates are to the ground truth. Motivated by ISI, the consistency of the components from one run to another can be measured using a global metric, cross-ISI, which we define as  $\text{ISI}_{ij}^c = \text{ISI}(\mathbf{P}^{ij})$ , where  $\mathbf{P}^{ij} = \mathbf{A}_i \mathbf{W}_j$  with elements denoted as  $p_{nm}^{ij}$ , where  $\mathbf{A}_i = \mathbf{W}_i^{-1}$  is the inverse of the demixing matrix of the  $i$ th run and  $\mathbf{W}_j$  is the demixing matrix of the  $j$ th run. To measure the consistency of a single run to all the other runs, the cross-ISI of the current run is generated by averaging all its pairwise cross-ISI:



**Fig. 1.** Scatter plot (left) and contour map (right) for cross-ISI and ISI pairs for 500 runs of each algorithm using GGD data sources. Also plotted are consistent runs selected by m-ICASSO (orange), MST (cyan) and cross-ISI (magenta).

$$ISI_i^C = \frac{1}{R-1} \sum_{j=1, j \neq i}^R ISI_{ij}^C \quad (4)$$

The calculation of cross-ISI is an algebraic function of the demixing matrix estimated from ICA, which makes it an objective measure and very efficient computationally.

### 3. EXPERIMENT AND NUMERICAL RESULTS

#### 3.1. Simulation set up

For the first set of experiments, simulated sources, each of which is distributed according to a GGD, are generated. GGD is a unimodal distribution and the shape parameter, denoted as  $\beta$  controls the peakedness and spread of the distribution. We use the definition in [23], such that when  $0 < \beta < 1$ , the distribution is super-Gaussian and a better match for fMRI sources. This selection also makes sure that Infomax, which uses a fixed nonlinearity that implies a specific super-Gaussian density model, can handle it. In our experiments, we study this range and here report results for two representative values,  $\beta = 0.2$  and  $0.6$ . In each scenario with the given shape parameter, 10 sources are generated such that  $\mathbf{S}_G \in \mathbb{R}^{10 \times 10^4}$  with 10000 samples. The mixing matrix  $\mathbf{A} \in \mathbb{R}^{10 \times 10}$  is randomly formed and selected as well-conditioned, such that the condition number is between 10 and 40.

For the second set of experiments, we use fMRI-like data using the SimTB toolbox [24, 25], the toolbox employs different parameters to control the generation of  $100 \times 100$  2D fMRI-like spatial maps and their corresponding time courses. A total of 10 spatial maps are selected as original sources ( $\mathbf{S}_s \in \mathbb{R}^{10 \times 10^4}$ ), and the associated time courses ( $\mathbf{T}_s \in \mathbb{R}^{260 \times 10}$ ) are generated, resulting in the mixture  $\mathbf{X}_s \in \mathbb{R}^{260 \times 10^4}$ . In this experiment, only one realization is used, and no noise is added to the data  $\mathbf{X}_s$ . The first step in processing the fMRI-like data consists of the application of principal component analysis (PCA). Since 10 sources are generated, the dimension of  $\mathbf{X}_s$  is reduced from 260 to 10.

Two parameters of SparseICA-EBM, sparsity parameter,  $\lambda$ , and smoothing parameter,  $\epsilon$ , are set for each dataset separately by making sure that the weight between the sparsity and independence matches the source model. For the GGD experiment, parameters are set as  $\lambda = 10^2$  and  $\epsilon = 10^{-2}$  and for SimTB generated data, parameters are set as  $\lambda = 10^4$  and  $\epsilon = 10^2$ . Estimation accuracy is measured by ISI defined in (3).

#### 3.2. Simulated data

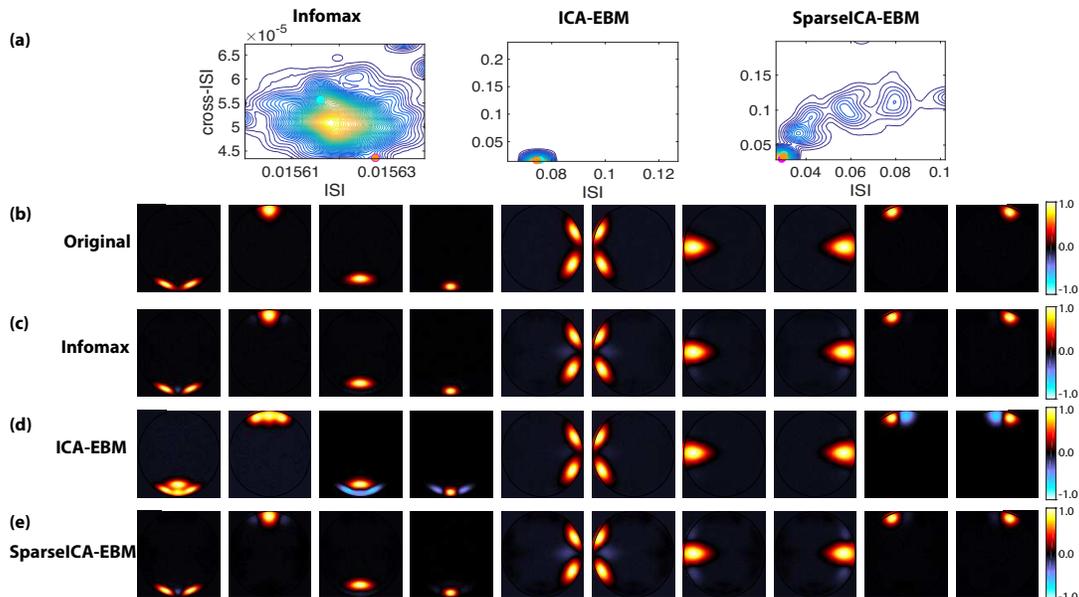
Each of the three ICA algorithms, Infomax, ICA-EBM, SparseICA-EBM, is run  $R=500$  times for the GGD experiment. Infomax uses the hyperbolic tangent function which implies a PDF that is proportional to  $\text{sech}(x)$  for the underlying sources. It can be shown that for sources that come from a GGD,  $\beta = 0.6$  is the closest to the PDF implied by the Infomax nonlinearity when values of  $\beta$  are investigated in increments of 0.1. The scatter plot and contour map of the

cross-ISI and ISI pairs for the 500 runs are displayed in Fig. 1. We see that solution from Infomax is more compact than ICA-EBM as it uses a fixed nonlinearity, and its bias is smaller when  $\beta = 0.6$ , as for this value, its nonlinearity is a good match to the source distribution. However, when  $\beta = 0.2$ , Infomax performs worst in terms of ISI due to the bias introduced as discussed in (2). The solution of SparseICA-EBM is also more compact than ICA-EBM since the solution space of SparseICA-EBM is shrunk compared to ICA-EBM. When  $\beta = 0.6$ , SparseICA-EBM performs worst since the sources are less sparse and its performance is desirable for  $\beta = 0.2$ . Additionally, we verified that there is a tight linear relationship between the cross-ISI and average cross correlation across estimates—global vs pairwise measures of consistency. Average cross correlation calculation is similar to that of cross-ISI but is more costly as it requires alignment of components.

For Infomax and SparseICA-EBM, the scatter plots show that all the solutions are close to each other providing similar accuracy. Even for ICA-EBM, whose solution space is relatively larger, the consistent runs are good representation of the average accuracy level of the algorithm. From the contour plots, we can see that the run with lowest ISI is usually far away from the run that is the most consistent, which means that it is not a good representation of the average performance of an algorithm. Consequently, it is less likely to be reproducible. This suggests that instead of finding the most accurate run—which is not an easy task in applications to real data where there is no ground truth—finding a run that is most consistent provides satisfactory performance in this case, especially when the overall model match is good.

In Fig. 1, the final runs selected from m-ICASSO, MST and cross-ISI are denoted by orange, cyan and magenta, respectively. Note that for cross-ISI, a lower value implies greater consistency. It shows that MST sometimes finds a run that is far away from the most consistent one, such as the case for SparseICA-EBM. In most of the scenarios, ISI of the consistent runs is close to the average. This demonstrates that the most consistent run is a good representation of the average algorithmic accuracy. Additionally, when model matches the true source PDF, the accuracy level of the consistent runs selected by three methods are close to each other. Sometimes cross-ISI yields a little higher accuracy while m-ICASSO yields higher accuracy in some of the cases. On the other hand, when there is a model mismatch, *i.e.*,  $\beta = 0.2$  for Infomax and  $\beta = 0.6$  for SparseICA-EBM, cross-ISI stands out and yields the most reasonable run. Additionally, for SparseICA-EBM (whose solution space is smaller, while also containing outliers), cross-ISI, which directly emphasizes the consistency, performs very well.

Cross-ISI is also computationally inexpensive. Computational time as in this case for 500 runs for the GGD experiment for cross-ISI is 15.72 seconds, which is significantly less than 974.01 seconds



**Fig. 2.** Contour maps for cross ISI and ISI for 500 runs of each algorithm and the spatial maps of the run selected by cross-ISI. Also plotted are consistent runs selected by m-ICASSO (orange), MST (cyan) and cross-ISI (magenta).

Data	Computational time (s)		
	m-ICASSO	MST	cross-ISI
GGD	974.01	912.04	15.72
fMRI-like	$1.07 \times 10^3$	$1.10 \times 10^3$	35.98

**Table 1.** Comparison of the computational time used by each consistent run selection method.

for m-ICASSO and 912.04 seconds for MST, as shown in Table. 1. It illustrates the potential of cross-ISI to accelerate the analysis when applied to real fMRI data or other problems.

### 3.3. fMRI-like data

Ten original components selected from SimTB toolbox are shown in Fig. 2(b), including fronto-parietal (F-P), default mode network (DMN), motor, frontal and visual components. Similar to the first set of experiments, we perform 500 runs on the mixture of these sources using each algorithm, and stable run selection is used to identify the most consistent run. Fig. 2(a) shows that the estimation accuracy of consistent runs selected by three algorithms is no less than the average level for ICA-EBM and SparseICA-EBM. For Infomax, the ISI of the consistent run selected by m-ICASSO and cross-ISI is a little higher than the average. This is somewhat acceptable, as the solution space of Infomax is very tight. As a measure of the independence of estimated components, we also evaluate the MI reduction (MIR) between the recovered components and the mixture, and mean remaining pairwise MI (PMI) between pairs of the components [4]. Results show that for ICA-EBM and SparseICA-EBM, ISI shows linear relationship with both MIR and PMI suggesting use of these measures for real data when there is no ground truth. Additionally, the MIR of the selected consistent runs using all three methods are virtually the highest among all runs, and their PMI are nearly the lowest for the two ICA algorithms, ICA-EBM and SparseICA-EBM. Computational time of these three consistent run selection methods are shown in Table. 1. Again cross-ISI is much more efficient than the other two.

Fig. 2(c)-(e) show the spatial maps of the consistent run with the lowest cross-ISI for three ICA algorithms. Components of the

run selected for Infomax and SparseICA-EBM are very close to the ground truth, with average spatial correlation values of 0.996 and 0.989, respectively. For the case of ICA-EBM, we observe that there are some components merged with each other, such as the first component with the third and fourth, the second with the ninth and tenth. From the PMI and pairwise correlation of the original 10 sources, we noted that the first, third and fourth components are not fully independent, as well as the second, ninth and tenth with higher PMI values compared to other pairs while the values of the other sources are much closer to zero. Hence, ICA-EBM seems to perform worst in terms of estimation accuracy as it is the only algorithm that maximizes independence by decreasing the bias in terms of KL divergence, as discussed in (2). Infomax and SparseICA-EBM on the other hand introduce a bias, which in this case enables a better model match as the sources are not fully independent but are all sparse. Due to space limitation we do not display the spatial maps selected by m-ICASSO and MST. However, they have been inspected and are similar to those shown in Fig. 2(c)-(e).

## 4. CONCLUSION

In this paper, trade-offs between the estimation accuracy and algorithmic consistency are explored in terms of model match. A new consistency measure, cross-ISI, is proposed to quantify algorithmic consistency and select a run among multiple ones in real world applications. Experimental results suggest that when modeling assumptions are met, the run selected by cross-ISI is a good representation of algorithmic accuracy. When we deviate from modeling assumptions, cross-ISI provides a more reliable selection than both m-ICASSO and MST, both widely used in fMRI analyses, making cross-ISI an attractive metric for this application among others. Furthermore, calculation of cross-ISI is an algebraic function of the demixing matrix estimated from ICA, making it suitable for a wide array of problems besides fMRI analysis and ideal for online and real-time applications, significantly reducing the cost of multiple runs. When working with real data, typically 10–20 independent runs are performed [13, 14] as both m-ICASSO and MST are computationally costly, with use of cross ISI this number can be significantly increased resulting in more reliable estimates for the subsequent analyses.

## 5. REFERENCES

- [1] T. Adalı, M. Anderson, and G.-S. Fu, "Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging," *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 18–33, 2014.
- [2] C. F. Beckmann, C. E. Mackay, N. Filippini, and S. M. Smith, "Group comparison of resting-state fMRI data using multi-subject ICA and dual regression," *NeuroImage*, vol. 47, no. Suppl 1, p. S148, 2009.
- [3] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.
- [4] A. Delorme, J. Palmer, J. Onton, R. Oostenveld, and S. Makeig, "Independent EEG sources are dipolar," *PLOS ONE*, vol. 7, no. 2, p. e30135, 2012.
- [5] D. T. Pham and P. Garat, "Blind separation of mixtures of independent sources through a quasi maximum likelihood approach," *IEEE Transactions on Signal Processing*, vol. 45, no. 7, pp. 1712–1725, 1997.
- [6] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic Press, 2010.
- [7] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1129–1159, 1995.
- [8] X.-L. Li and T. Adalı, "Independent component analysis by entropy bound minimization," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5151–5164, 2010.
- [9] J.-F. Cardoso, "Infomax and maximum likelihood for blind source separation," *IEEE Signal Processing Letters*, vol. 4, no. 4, pp. 112–114, 1997.
- [10] J. Himberg and A. Hyvarinen, "ICASSO: software for investigating the reliability of ICA estimates by clustering and visualization," in *Neural Networks for Signal Processing, 2003. NNSP'03. 2003 IEEE 13th Workshop on*, 2003, pp. 259–268.
- [11] E. S. Finn, X. Shen, D. Scheinost, M. D. Rosenberg, J. Huang, M. M. Chun, X. Papademetris, and R. T. Constable, "Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity," *Nature Neuroscience*, vol. 18, no. 11, pp. 1664–1671, 2015.
- [12] T. Adalı, Y. Levin-Schwartz, and V. D. Calhoun, "Multimodal data fusion using source separation: Application to medical imaging," *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1494–1506, Sep. 2015.
- [13] S. Ma, N. M. Correa, X.-L. Li, T. Eichele, V. D. Calhoun, and T. Adalı, "Automatic identification of functional clusters in fMRI data using spatial dependence," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 12, pp. 3406–3417, 2011.
- [14] W. Du, S. Ma, G.-S. Fu, V. D. Calhoun, and T. Adalı, "A novel approach for assessing reliability of ICA for fMRI analysis," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, 2014, pp. 2084–2088.
- [15] GIFT, "Group ICA of fMRI toolbox (GIFT)," 2011. [Online]. Available: <http://mialab.mrn.org/software/gift/>
- [16] X.-N. Zuo, C. Kelly, J. S. Adelstein, D. F. Klein, F. X. Castellanos, and M. P. Milham, "Reliable intrinsic connectivity networks: Test–retest evaluation using ICA and dual regression approach," *NeuroImage*, vol. 49, no. 3, pp. 2163–2177, 2010.
- [17] S. C. Strother, J. Anderson, L. K. Hansen, U. Kjems, R. Kustra, J. Sidtis, S. Frutiger, S. Muley, S. LaConte, and D. Rottenberg, "The quantitative evaluation of functional neuroimaging experiments: the NPAIRS data analysis framework," *NeuroImage*, vol. 15, no. 4, pp. 747–771, 2002.
- [18] M. N. Wernick, Y. Yang, J. G. Brankov, G. Yourganov, and S. C. Strother, "Machine learning in medical imaging," *IEEE Signal Processing Magazine*, vol. 27, no. 4, pp. 25–38, 2010.
- [19] Z. Boukouvalas, Y. Levin-Schwartz, and T. Adalı, "Enhancing ICA performance by exploiting sparsity: Application to fMRI analysis," in *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*, 2017, pp. 2532–2536.
- [20] W. Du, H. Li, X.-L. Li, V. D. Calhoun, and T. Adalı, "ICA of fMRI data: performance of three ICA algorithms and the importance of taking correlation information into account," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, 2011, pp. 1573–1576.
- [21] Z. Boukouvalas, Y. Levin-Schwartz, V. D. Calhoun, and T. Adalı, "Sparsity and independence: Balancing two objectives in optimization for source separation with application to fMRI analysis," *Journal of the Franklin Institute*, 2017.
- [22] W. Du, Y. Levin-Schwartz, G.-S. Fu, S. Ma, V. D. Calhoun, and T. Adalı, "The role of diversity in complex ICA algorithms for fMRI analysis," *Journal of Neuroscience Methods*, vol. 264, pp. 129–135, 2016.
- [23] E. Gómez, M. Gomez-Viilegas, and J. Marin, "A multivariate generalization of the power exponential family of distributions," *Communications in Statistics-Theory and Methods*, vol. 27, no. 3, pp. 589–600, 1998.
- [24] E. B. Erhardt, E. A. Allen, Y. Wei, T. Eichele, and V. D. Calhoun, "SimTB, a simulation toolbox for fMRI data under a model of spatiotemporal separability," *Neuroimage*, vol. 59, no. 4, pp. 4160–4167, 2012.
- [25] S. Ma, R. Phlypo, V. D. Calhoun, and T. Adalı, "Capturing group variability using IVA: a simulation study and graph-theoretical analysis," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, 2013, pp. 3128–3132.